ECE 693 – Special Topics:
AI for Radar System Design

Introduction to Machine Learning

Dr. Sevgi Zubeyde Gurbuz
szgurbuz@ua.edu

Jan. 28, 2022

THE UNIVERSITY OF ALABAMA®

What is Learning?

- (Arthur Samuel, 1959)
  - Field of study that gives computers the ability to learn without being explicitly programmed.
- (Tom Mitchel, 1998)
  - A computer program is said to learn from some experience E with respect to some task T and some performance measure P, if its performance T as measured by P, improves with experience E.

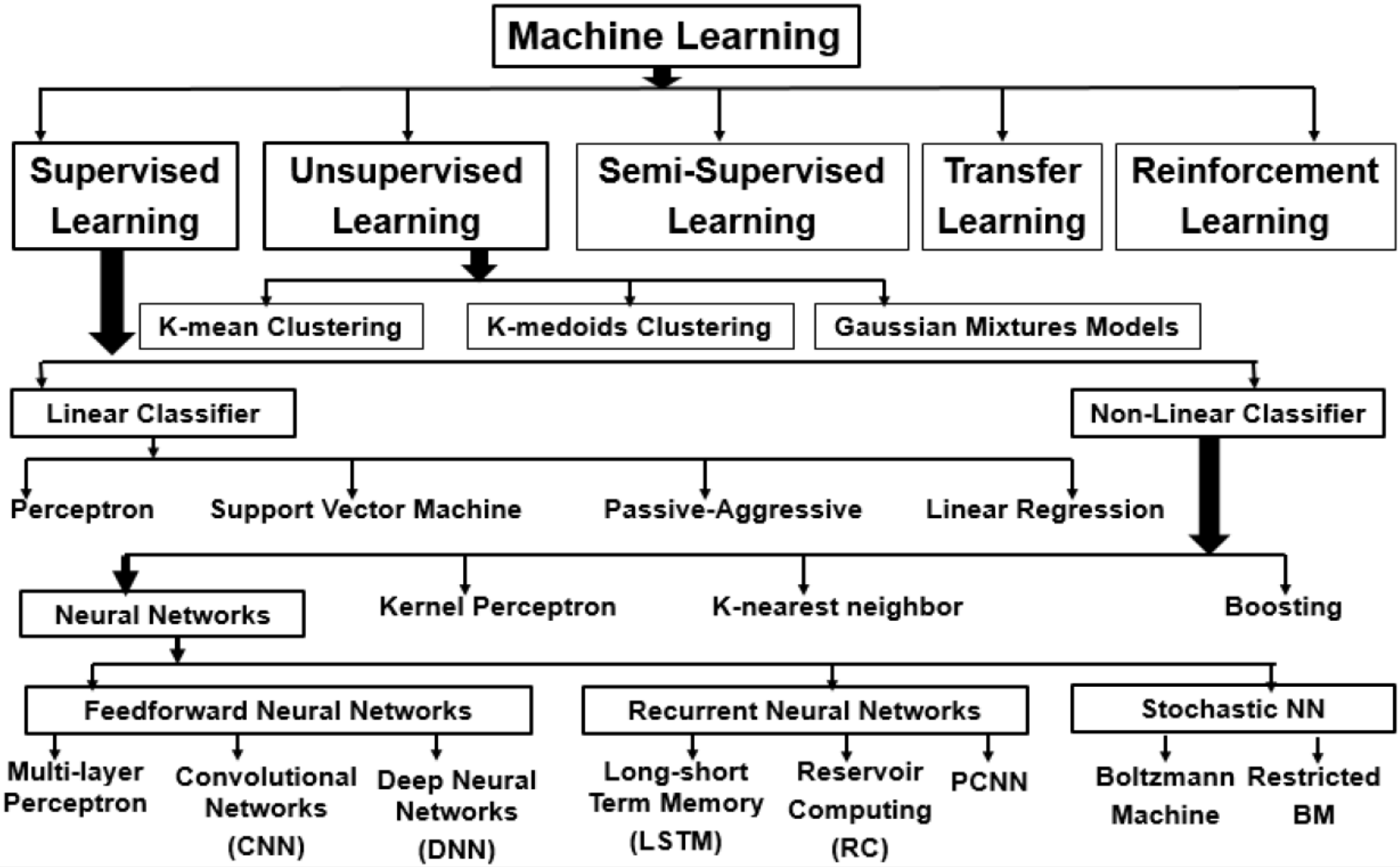
Example

- Suppose your computer watches which emails you do or do not mark as spam, and based on that learns how to better filter spam.

  What are task T, experience E, and performance P?

  - ❑ Classifying emails as spam or not spam.
  - ❑ Watching you label emails as spam or not spam.
  - ❑ The number of emails correctly classified as spam or not spam

Algorithm Taxonomy



Components of ML Algorithms

- Representation
  - Decision trees
  - Sets of rules / Logic programs
  - Graphical models (Bayes/Markov nets)
  - Neural networks
  - Support vector machine
- Evaluation
  - Accuracy, precision and recall, mean-squared error, likelihood, posterior probability, cost function, margin of error, entropy, etc.
- Optimization
  - Combinatorial (e.g. greedy search)
  - Convex (e.g. gradient descent)
  - Constrained (e.g. linear programming)

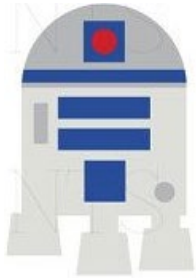
Types of Experiences

- Supervised Learning
  - Humans learn from past experiences
  - Computers learn from labeled data that represents past experience or prior knowledge
- Unsupervised Learning
  - No *a priori* models, unlabelled data
- Reinforcement Learning
  - Training through interaction with an environment; the agent receives a numerical reward signal

Unsupervised Learning

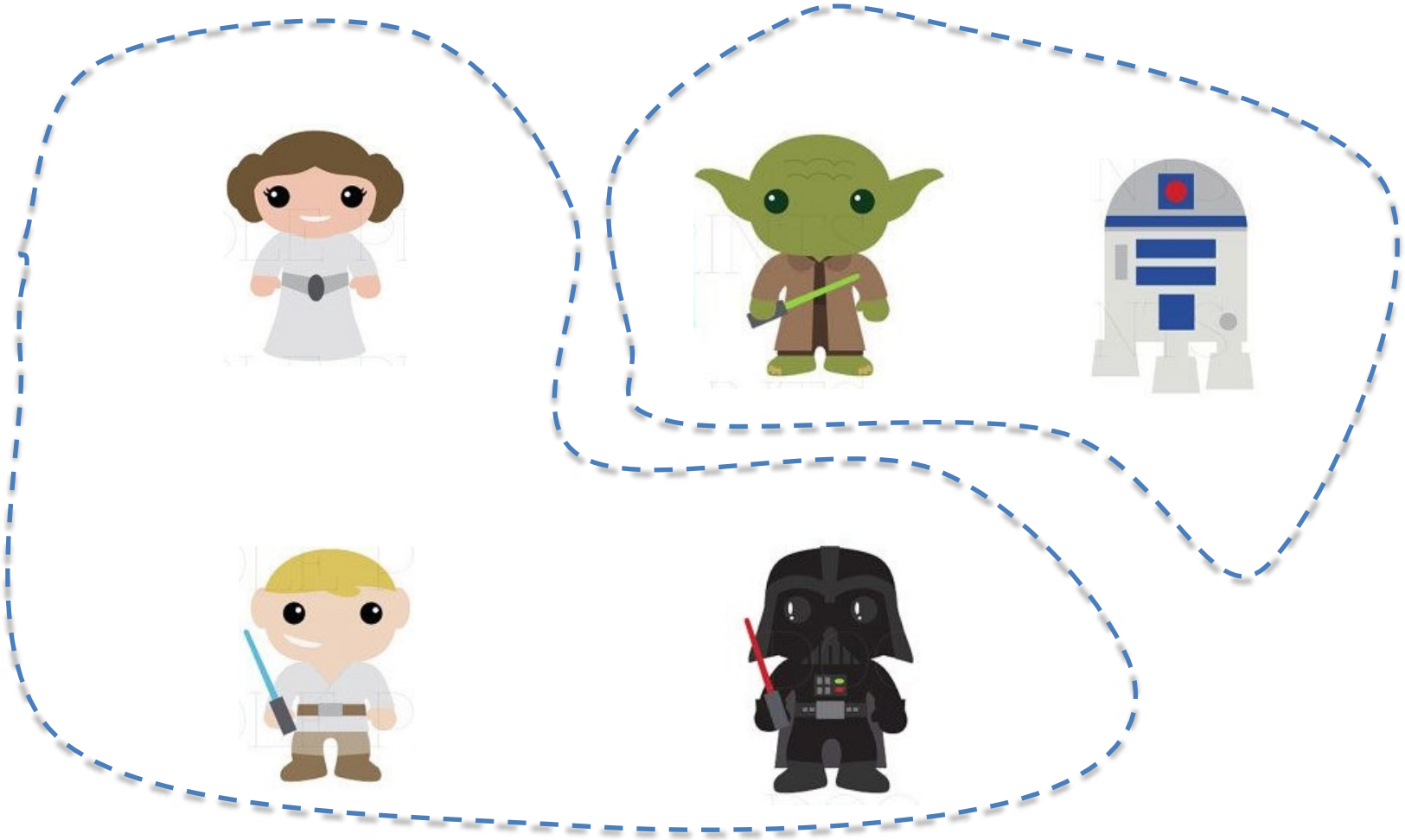
Challenge: Classify into two groups



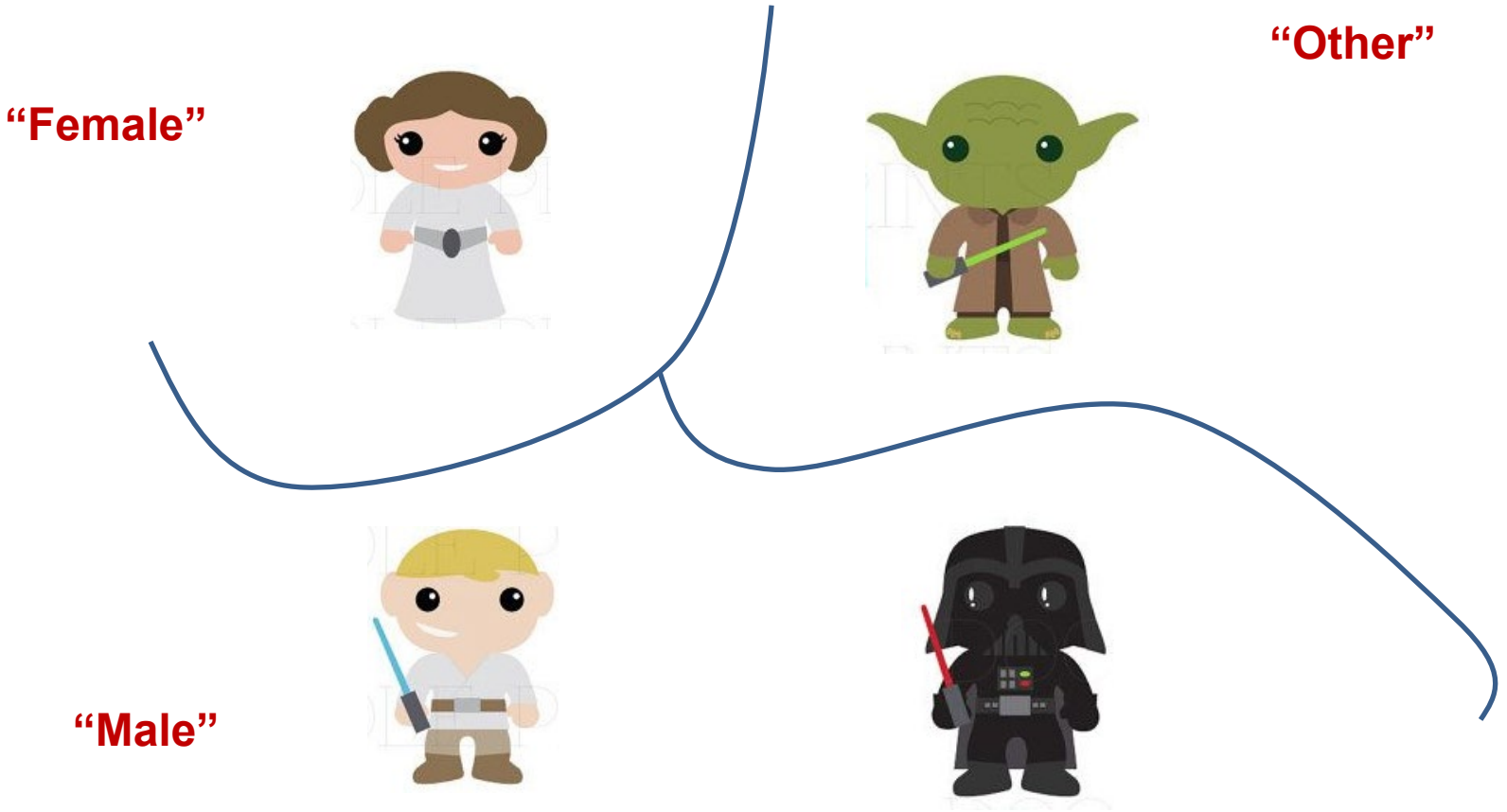


Sevgi Z. Gurbuz (szgurbuz@ua.edu)

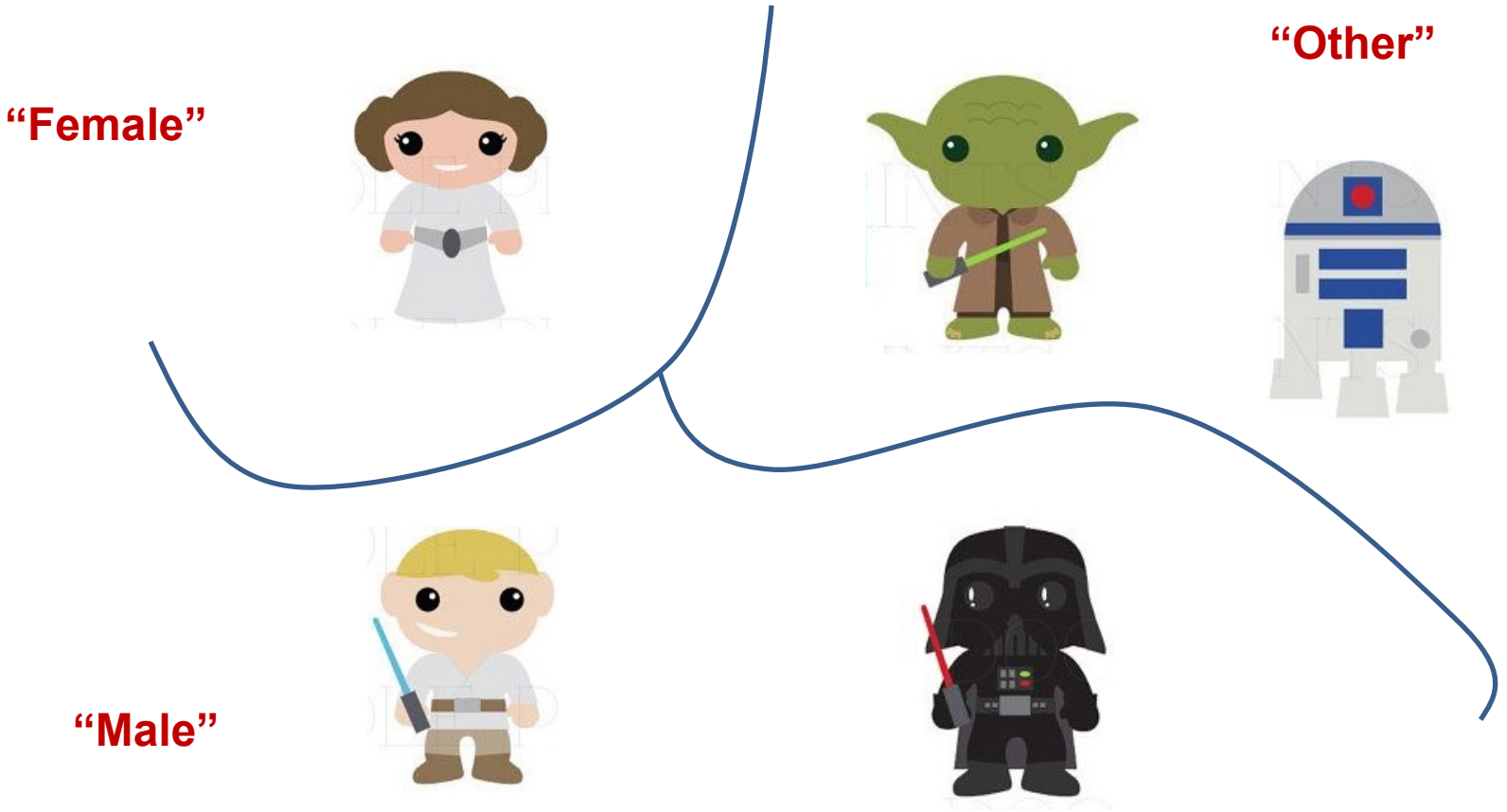
Unsupervised Learning: Human vs. Other



Supervised Learning: The Training Process

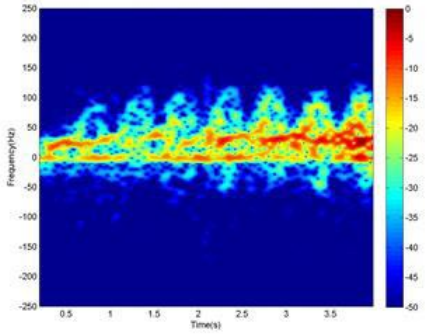


Supervised Learning: Testing

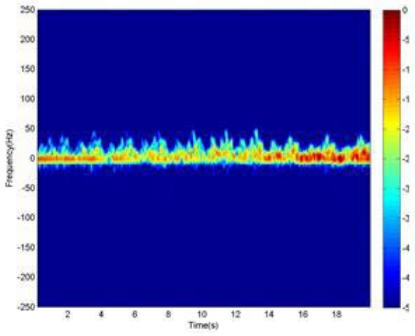


The Classification Process

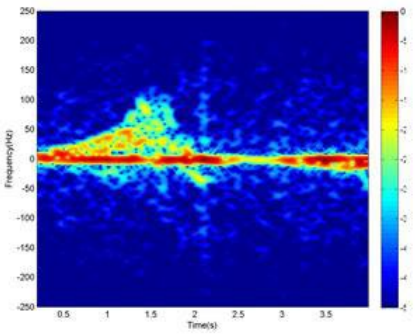
Walking



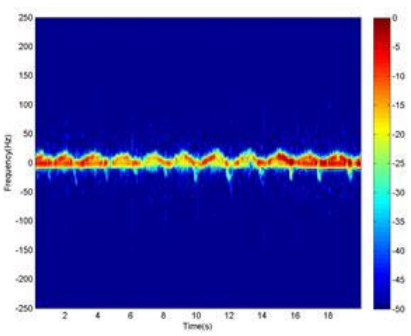
Limping



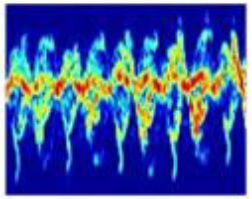
Falling



Wheelchair Use



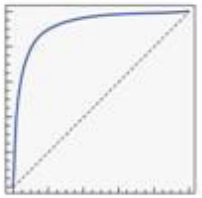
INPUT



CLASSIFICATION PROCESS

FEATURE EXTRACTION → FEATURE SELECTION → CLASSIFICATION

OUTPUT



- Raw I/Q data
- Spectrogram

- Physical
- Speech-Inspired
- Transform-Based

- Dimension Reduction
- Filters
- Wrappers

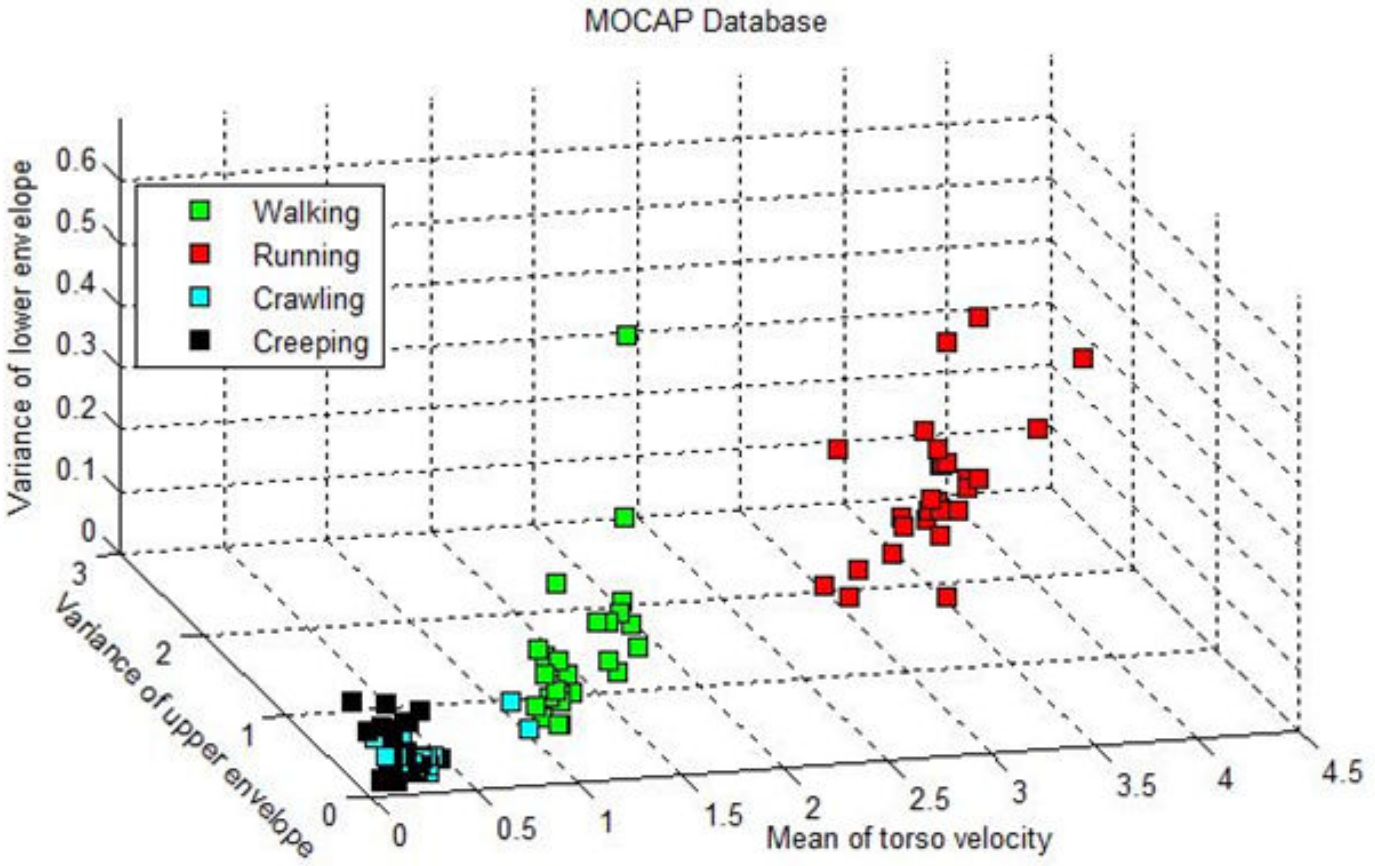
- SVM
- kNN
- Bayesian

...and more

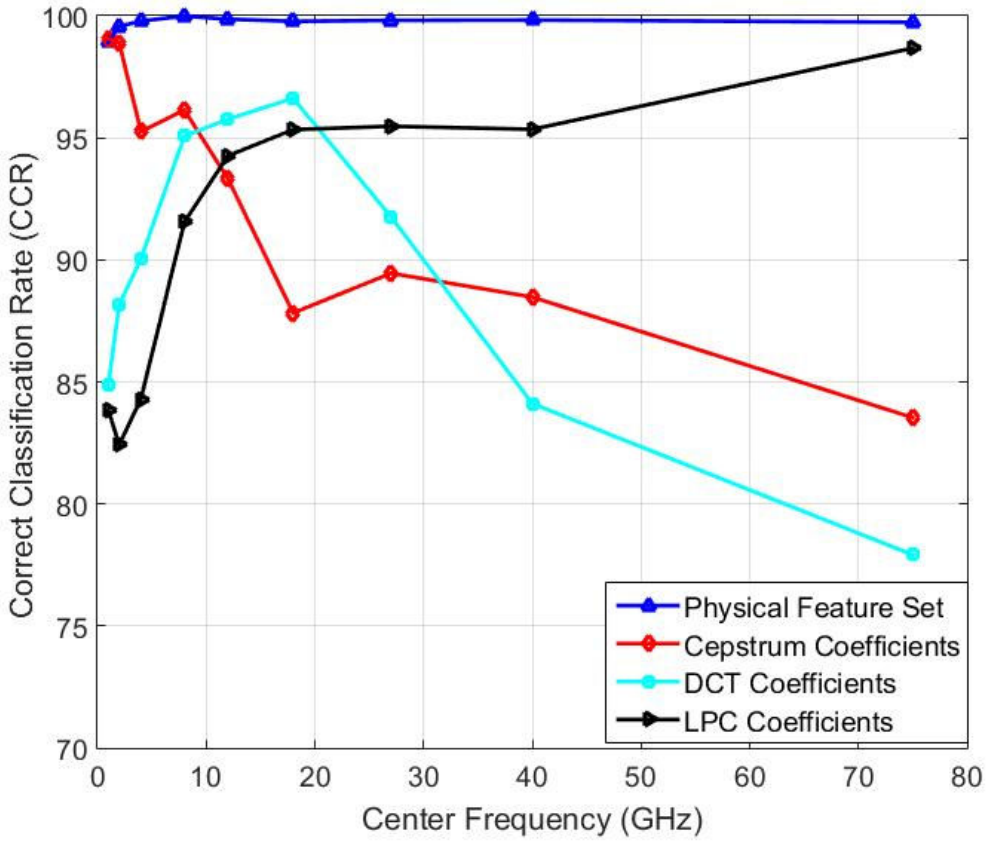
- ROC curve
- Confusion matrix

Sevgi Z. Gurbuz (szgurbuz@ua.edu)

What Features We Extract is Important



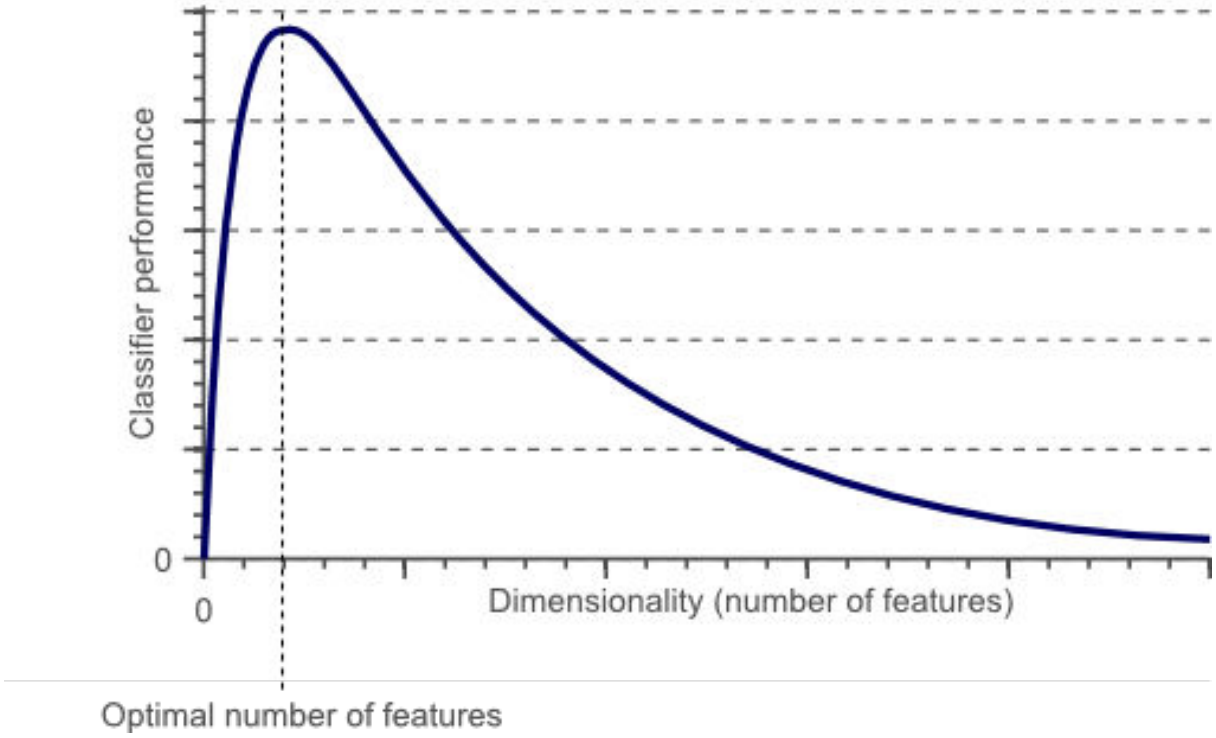
Efficacy of Features Dependent on Transmit Parameters!



S.Z. Gurbuz, B. Erol, B. Cagliyan, B. Tekeli, "Operational Assessment and Adaptive Selection of Micro-Doppler Features," IET RSN, 2015.

More Features Not Necessarily Better

- "Curse of Dimensionality"

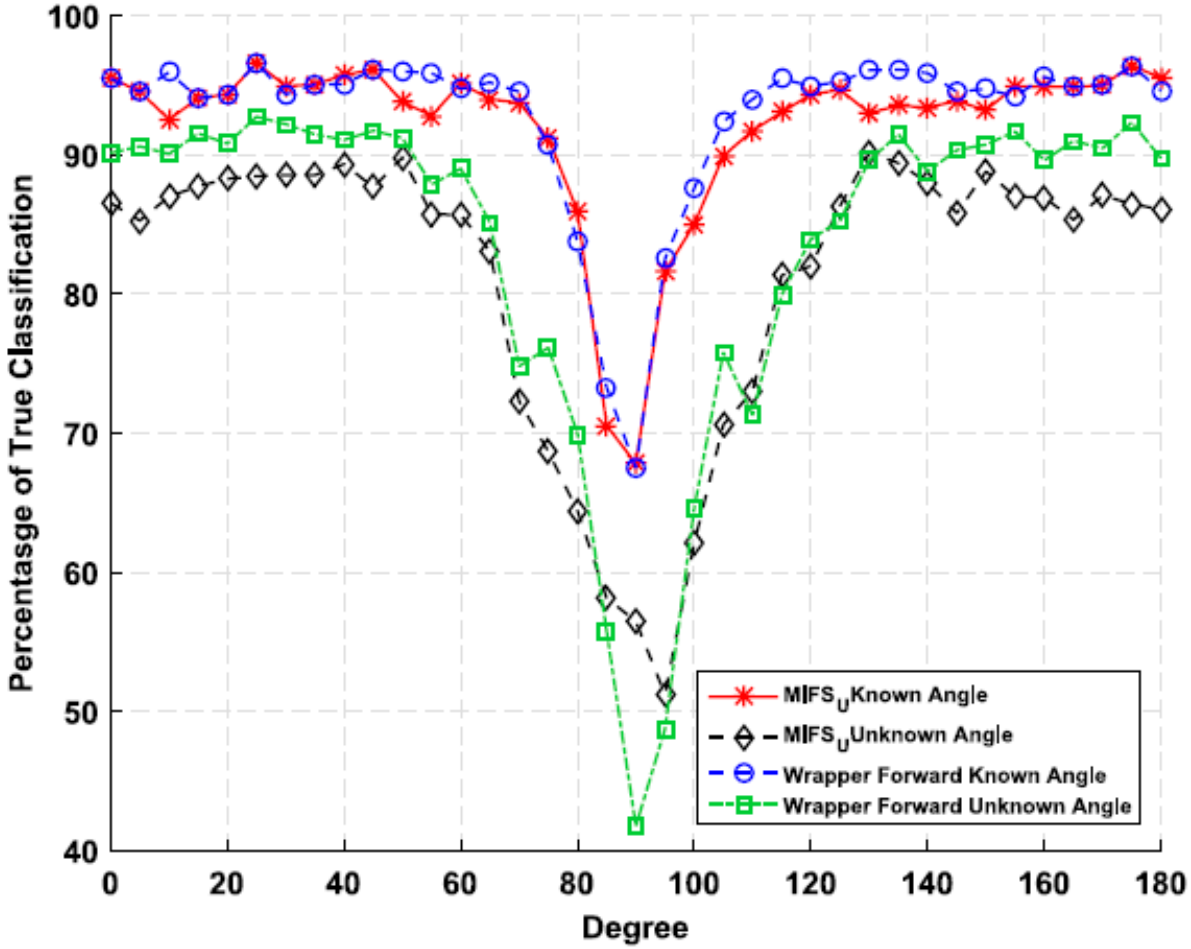


https://www.visiondummy.com/2014/04/curse-dimensionality-affect-classification/

Feature Selection Algorithms

- Filter Methods
  - Based on some metric, features are selected
  - Advantage is that it doesn't depend on classifier
  - Disadvantage is that it may not accurately predict ultimate classification accuracy
- Wrapper Methods
  - Brute force just try and find the best combination of features
  - May (or may not) give better combination, but computationally intensive and classifier specific

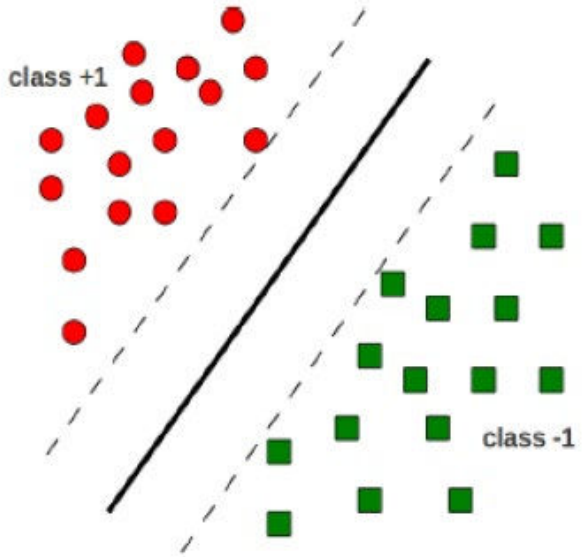
Feature Selection Example



Sevgi Z. Gurbuz (szgurbuz@ua.edu)

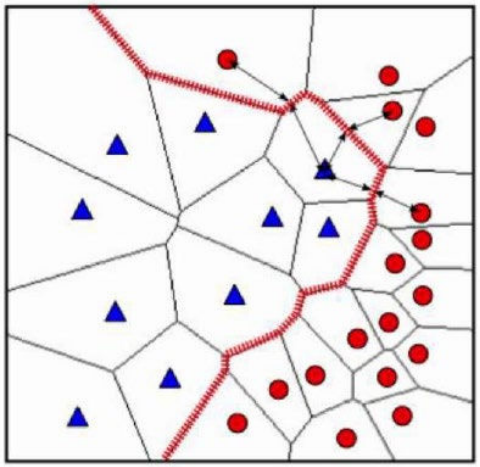
Classifiers in Supervised Learning

- Linear Classifier: Support Vector Machine (SVM)
  - Aims to maximize the margin (**distance**) between two classes in the input space



Non-Linear Classifier Example

- K-Nearest Neighbors
  - Algorithm searches the entire training set, calculating the difference between the new sample and each training model
  - Output is the class with the K-most similar neighbors



Other Classifiers

Supervised:

- Decision Trees
- Random Forests
- Bayesian Classifiers
- Boosting
- Neural Networks

… and more…

Unsupervised:

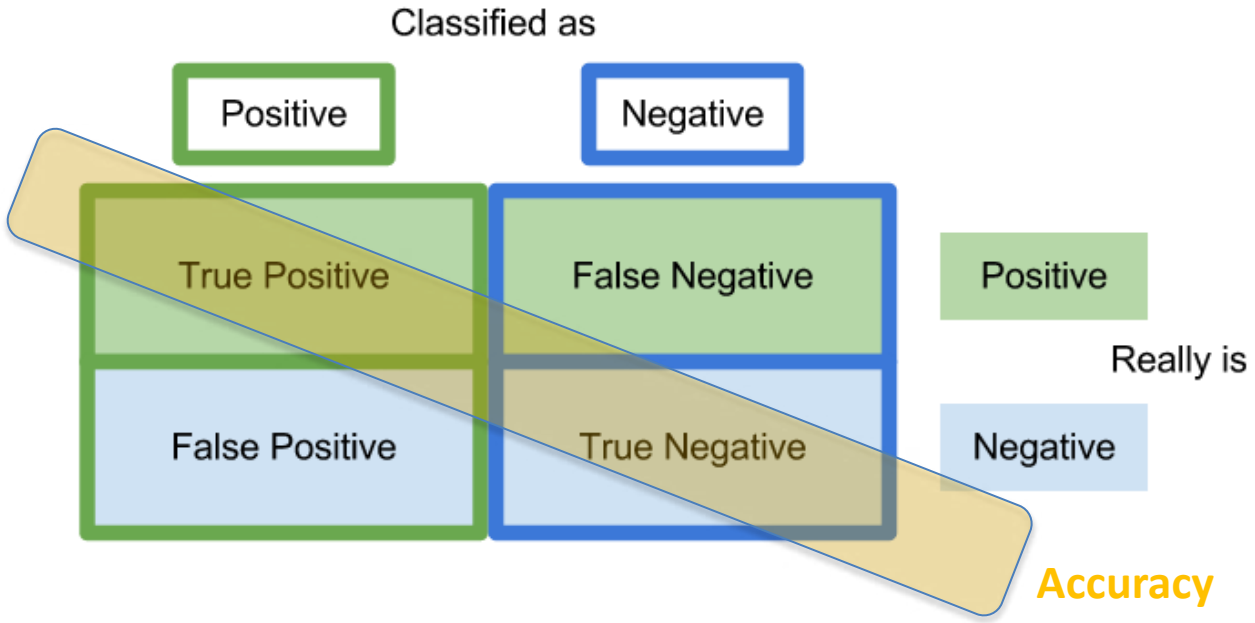
- Clustering
- Gaussian Mixture Models

… and more…

Distance Metrics

Performance Metrics

- Confusion Matrix



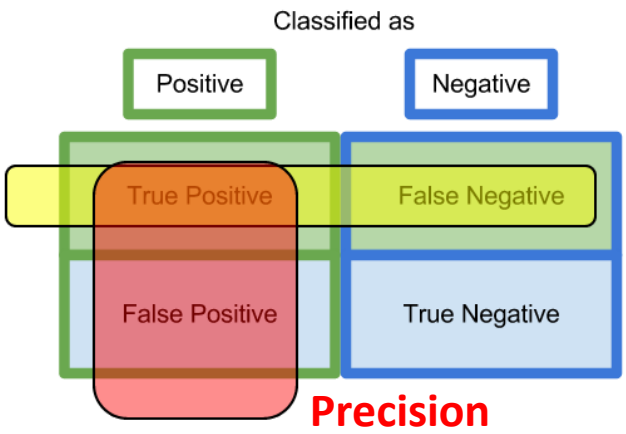
- Accuracy
  - Correct classification rate (# correct/total)

More Metrics

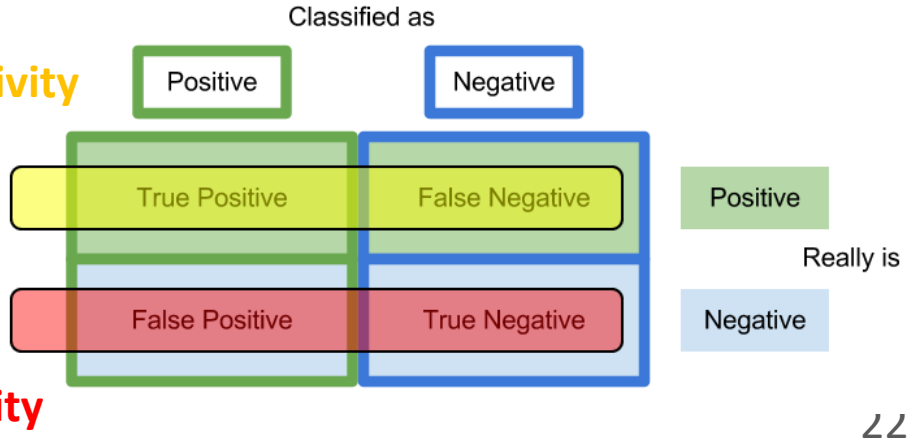
- Precision = true positives / all positives
- Recall = Sensitivity =

  = true positives/(true positive + false negative)
- Specificity =

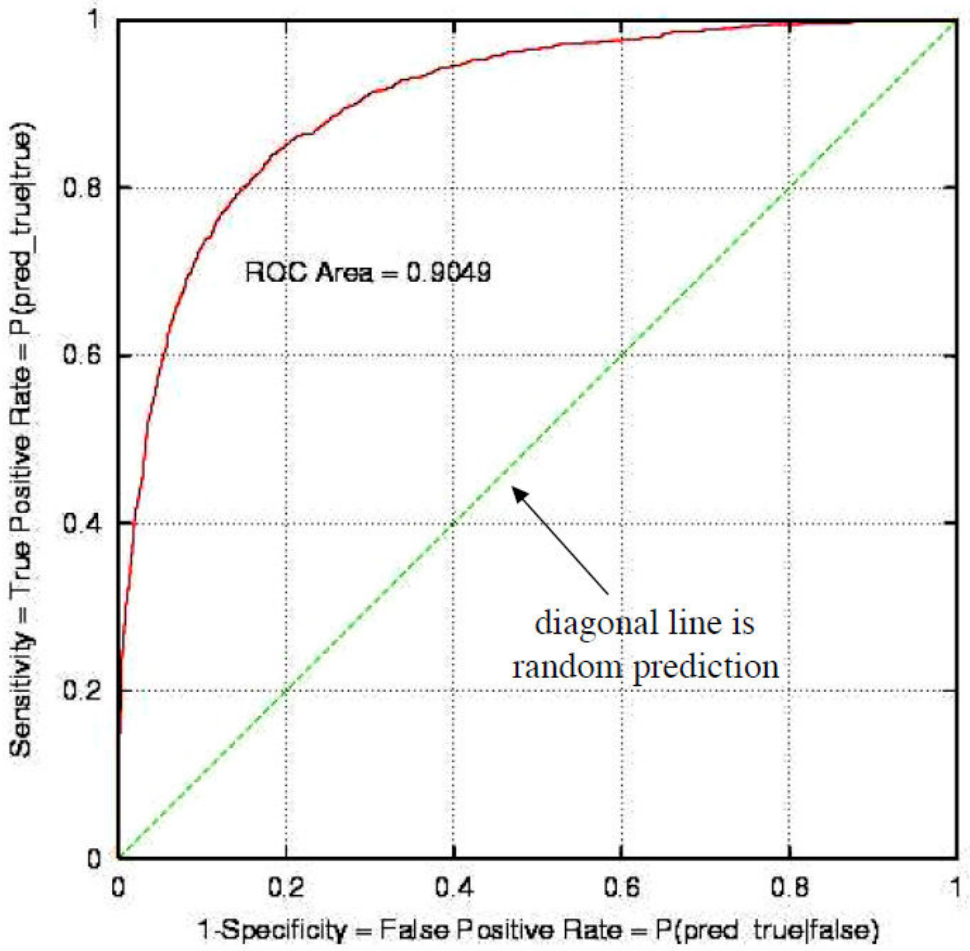
  = true negatives/(true negatives + false positives)



Recall = Sensitivity



Receiver Operating Characteristics (ROC) and Area-Under-Curve (AUC)



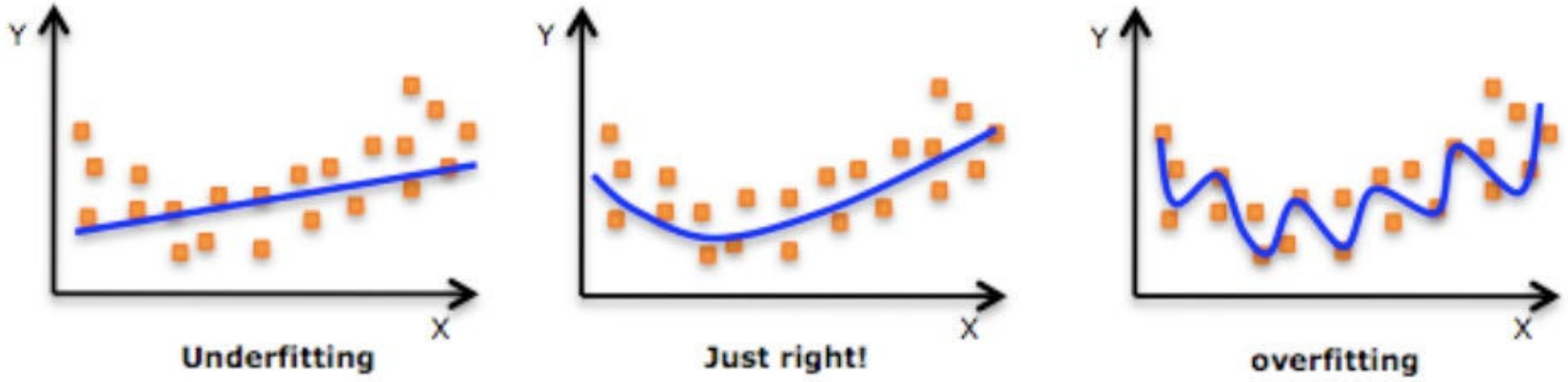
Watch Out for Generalization and Overfitting!

- Generalization
  - The ability to produce reasonable outputs for inputs not encountered during training

In other words: NO PANIC when "never seen before" data are given in input!

Overfitting

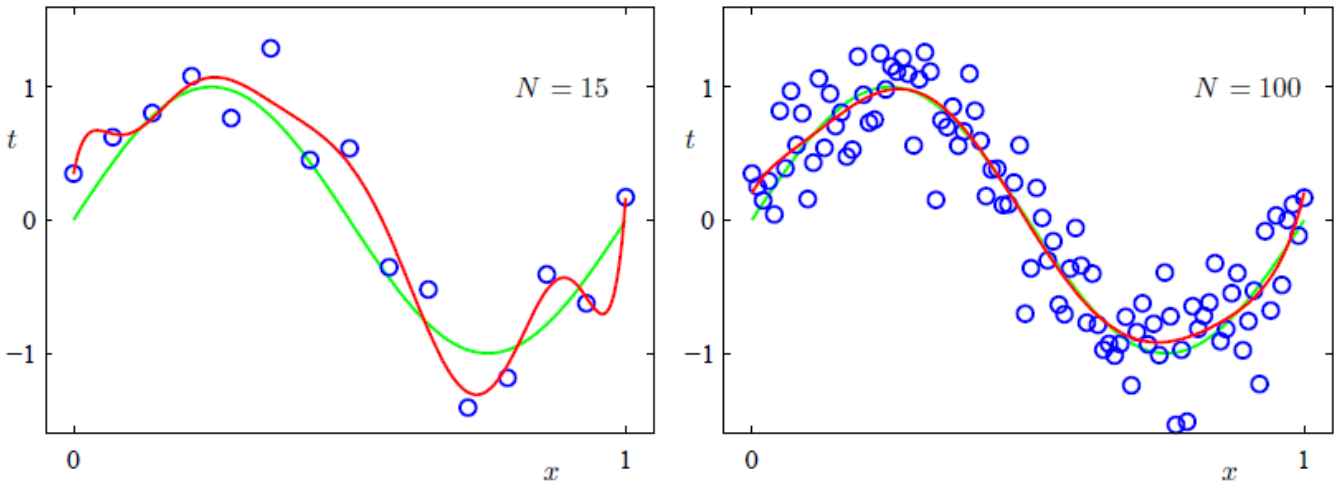
- Overfitting is a modeling error that introduces bias to the model because it is too closely related to the data set.
- Overfitting makes the model relevant to its data set only, and irrelevant to any other data sets.
- Performs well on the data used during training but poorly with new data



To avoid overfitting, use a sufficient amount of training data with statistical variations, and early stopping...

The More (Data), The Merrier

- Overfitting depends on the amount of data, relative to the complexity of the hypothesis
- With more data, we can explore more complex hypotheses spaces, and still find a good solution

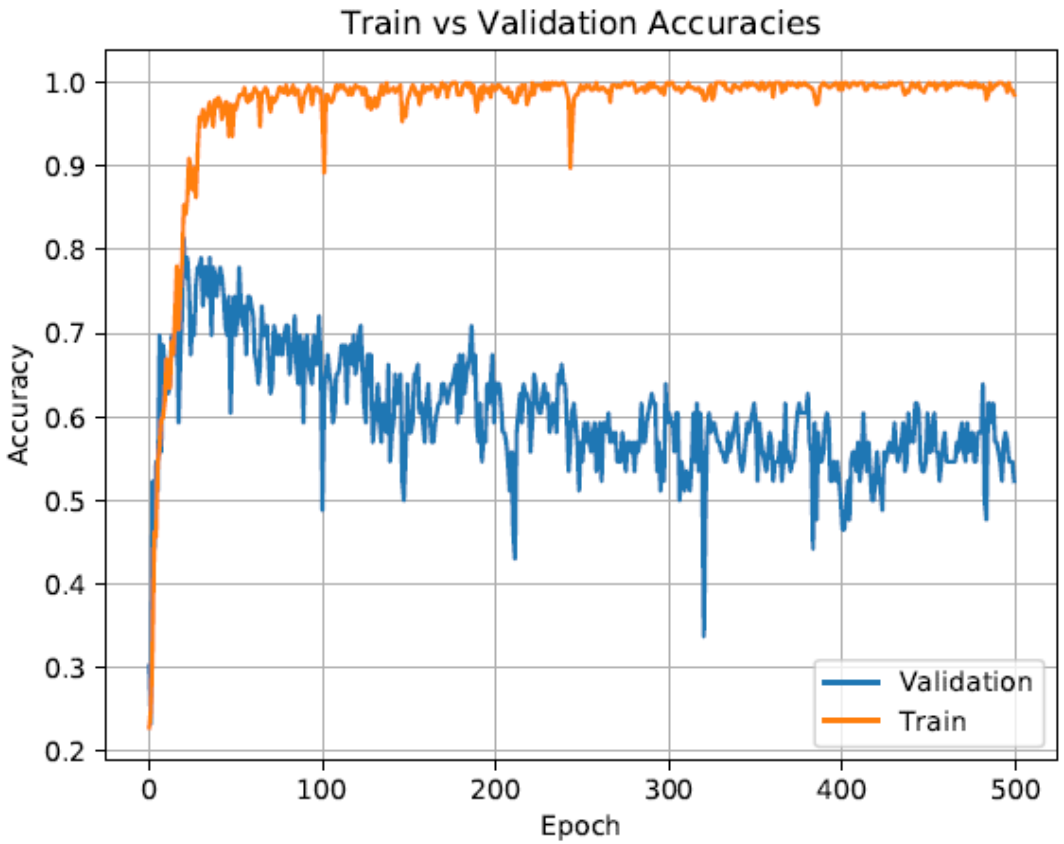


How Can You Detect Overfitting?

- Divide data into three disjoint sets:
  - Training: set of examples for learning
  - Validation: examples to tune architecture

    of assess errors
  - Test: performance assessment of classifier; gives unbiased estimate of generalization error
- Overfitting can be identified by checking validation metrics such as accuracy and loss. The validation metrics usually increase until a point where they stagnate or start declining when the model is affected by overfitting.

Overfitting Example



How Can Overfitting Be Prevented?

- Training with More Data
- Data Augmentation
- Reduce Complexity of Model
- Ensembling (combine predictions from two or more separate models)